Computing resilience of interconnected systems by piecewise linear Lyapunov functions

Alberto Tacchella and Armando Tacchella

KEYWORDS

Simulation of Control Systems, Piecewise-Linear Switched Systems, Cyber-Security and Critical Infrastructure Protection

ABSTRACT

Resilience, i.e., the ability to withstand and recover from disruption, is a fundamental requirement for mission-critical automation systems. Interconnection among systems makes the analytical determination of resilience zones harder than in isolated systems, because the failure of a system usually has an impact on those connected to it. In this paper we propose an algorithm to determine resilience zones of interconnected systems based on the computation of piecewise linear Lyapunov functions. The algorithm is based on a model introduced previously that abstracts from specific system dynamics and enables analysis in the space of performances. Experiments with an interconnected system inspired to a real wastewater treatment facility show the feasibility and the accountability of our approach.

I. INTRODUCTION

The number of security incidents affecting industrial automation systems has been steadily increasing over the past few years — see, e.g., [Lou15]. The main problem is that more and more such systems do not work in isolation, but are routinely connected between them, often relying on wide-area networks including the Internet. Through such connections, cyber-attacks can be brought to the systems and cause disruption in services, damage to equipment or severe impairment of human activities. Detecting weaknesses, fixing them and monitoring critical events in industrial automation systems are compelling and heavily investigated matters, but we must also acknowledge that, in spite of all the efforts made to secure them, connected systems may never be fully secure. In this scenario, the concept of *resilience* emerges as an additional target, complementary to prevention and protection from attacks, but not less important. This line of thought is pervasive in the Presidential Policy Directive 21 [Oba13] about the security of critical infrastructure, which defines resilience as *"[...] the ability to [...] withstand and recover rapidly from disruptions. Resilience includes the ability to withstand and recover from deliberate attacks, accidents, or naturally occurring threats or incidents"*. More recently, the term *cyber-resilience* has been coined to identify specifically *"the ability to continuously deliver the intended outcome despite adverse cyber events"* [BHSZ15] and this is the interpretation whereto we adhere in the following.

As mentioned in [AF13], the analysis of interconnected automation systems involves a number of challenging issues. One such issue is that they are built as hierarchies of subsystems which makes their design, implementation and maintenance feasible from an engineeering point of view. However, the analysis based on decomposition in systems components is either hardly feasible or of poor utility. If we consider resilience, one problem is the difficulty of inferring the resilience of the overall system considering the resilience of its components. The well-known phenomenon of *cascading failures* — see, e.g., [CLM04] — is an example of how individually healthy subsystems might become unstable and fail because other connected subsystems cease to function properly, leading to failures that potentially affect the whole system. In case of resilience it is important to understand how variations in the behavior of one system can affect the others, and in particular whether outside the nominal operating range there exist *resilience regions*, i.e., zones in the system state space wherein a failure of some component can be "absorbed" by the system which remains functional, albeit possibly at a degraded level.

The problem of devising models to compute resilience in interconnected systems has been studied previously. In [AF13], the authors introduce a model where systems are treated as abstract entities, decomposed into a set of elements interconnected by functional dependencies. The failures considered are of two types: internal operation drifts and external cascade effects. The modeling approach is quite general, yet simple enough to be applied on a variety of systems. While we use the contribution of [AF13] as a basis for ours, the topology analized in that paper applies only to small systems consisting of two components connected through a "feedback" loop. A slight generalization of the results in [AF13] is presented in [LPZ14] where the authors consider also "loop structures" for the analytical determination of resilience. Finally, in [LFZ17], the application of the framework introduced in [AF13] is considered in the context of critical infrastructure.

In this paper we contribute an extension of the methodology proposed in [AF13] which in principle allows to infer automatically Lyapunov functions to study the stability, and thus the resilience of interconnected systems. To the extent of our knowledge, this is the first time that a computational approach is inves-

Alberto Tacchella is with Fondazione Bruno Kessler, Trento, Italy — E-mail: atacchella@fbk.eu. Armando Tacchella is with DIBRIS, University of Genoa, Genoa, Italy — E-mail: armando.tacchella@unige.it

Communications of the ECMS, Volume 34, Issue 1,
Proceedings, ©ECMS Mike Steglich, Christian Mueller,
Gaby Neumann, Mathias Walther (Editors)
ISBN: 978-3-937436-68-5/978-3-937436-69-2(CD) ISSN 2522-2414

tigated to attack this problem, providing also some experimental evidence of its feasibility and accountability on a case study extracted from a real critical infrastructure. More in detail, our contribution can be fleshed out as follows:

- An iterative method to compute piecewise linear Lyapunov functions characterizing the resilience of specific topologies of interconnected systems;
- A case study about the model of an urban wastewater treatment facility;
- Experimental results on the case study showing promise about the method.

The rest of the paper is structured as follows. In Section II we introduce basic terminology, notation and definitions related to Lyapunov stability, piecewise linear switching systems, and the resilience model for interconnected systems that we consider as a basis for our investigation. In Section III we extend the basic resilience model to systems having a specific topology and, for systems having such topology, in Section IV we present an approach to compute piecewise linear Lyapunov functions to tackle their stability. Finally, in Section V, we introduce our case study related to wastewater treatment, show its simulation model, how we extracted relevant parameters from it, and the results of applying our approach to the system under consideration. We conclude in Section VI with some final remarks and an agenda for future developments.

II. BACKGROUND

We recall briefly the method of Lyapunov functions for proving the stability of an equilibrium point of a nonlinear system; see e.g. Khalil [Kha92] for a textbook introduction. We also refer the reader to [GH15] for a comprehensive review on Lyapunov functions and related computational methods.

Let us consider an autonomous, first-order system of ordinary differential equations (ODEs)

$$\dot{\mathbf{x}} = f(\mathbf{x}),$$

where f is a C^1 function $\mathbb{R}^n \to \mathbb{R}^n$. We assume that the corresponding dynamical system has (at least) an equilibrium point, which can be taken to be at $0 \in \mathbb{R}^n$ without loss of generality. We denote by $\Phi_t(\mathbf{x}_0)$ the corresponding flow at time t with initial value $\mathbf{x}_0$.

The equilibrium point at 0 is called:

- *stable* if for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $\mathbf{x}_0$ with $\|\mathbf{x}_0\| < \delta$ we have that $\|\Phi_t(\mathbf{x}_0)\| < \varepsilon$ for every $t \geq 0$;
- *asymptotically stable* if it is stable and there exists a $\delta' > 0$ such that for every $\mathbf{x}_0$ with $\|\mathbf{x}_0\| < \delta'$ we have that $\lim_{t\to\infty}\|\Phi_t(\mathbf{x}_0)\| = 0$;
- *exponentially stable* if it is asymptotically stable and the rate of convergence to zero of $\|\Phi_t(\mathbf{x}_0)\|$ is exponential.

Classically, a *Lyapunov function* for the system (at the given equilibrium point) is a C^1 function $V \colon U \to \mathbb{R}$ defined on a neighborhood U of 0 and such that:

- V is positive away from 0 and $V(0) = 0$ (in other words, V is a positive function with a global minimum at 0), and
- V is strictly decreasing along the solution curves of the ODE (outside of the equilibrium point). If we denote by $\dot{V}(\mathbf{x})$ the derivative of V along the orbit passing through $\mathbf{x}$, then a sufficient condition for this to happen is that $\dot{V}(\mathbf{x}) < 0$ for every $\mathbf{x} \in U \setminus \{0\}$.

Lyapunov's theorem asserts that the existence of a Lyapunov function is sufficient to guarantee the asymptotic stability of the equilibrium. Moreover, any sublevel set of V which is contained in U is also a subset of the basin of attraction of the equilibrium point.

For the purposes of the present paper it is important to note that the above formulation of Lyapunov's theorem can be considerably weakened, for instance by removing the hypothesis that the Lyapunov function be C^1 or even continuous. The only important requirement is that V must decrease in a suitable sense along the trajectories of the system. In the sequel we shall always speak of Lyapunov functions in this more general sense.

A *switched system* is a hybrid system in which a continuous-time evolution law is coupled to a set of discrete "switching" events. Typically, the continuous state space is partitioned into a finite number of *operating regions* by means of a family of *switching surfaces*. In each operating region, the evolution of the system is described by a given system of ODEs. Whenever the trajectory of the system hits a switching surface, the continuous state jumps instantaneously to a new value, specified by a *reset map*. When the reset map is trivial, that is the switching events only involve a change in the continuous evolution law, one speaks of an *autonomous* switched system, or *switching system*. For a general introduction to switched (and switching) systems we refer the reader to the books by Liberzon [Lib03], or Sun and Ge [SG11].

We are interested in a particular subclass of switching systems, namely the ones in which the state space is a finite-dimensional linear space (say $\mathbb{R}^n$ for some $n \in \mathbb{N}$) and in each operating region the continuous evolution law is either linear or affine. Following Johansson [Joh03], we shall call such systems *piecewise linear* (PWL) *switching systems*[1].

A piecewise linear switching system is then defined by:

- a partition $\{X_i\}_{i\in I}$ of the state space $\mathbb{R}^n$ in a set of operating regions, and
- for each $i \in I$ an affine evolution law of the form

$$\dot{\mathbf{x}} = A_i\mathbf{x} + b_i$$

where A_i is a $n \times n$ matrix and b_i is a column vector.

Let us recall from [Joh03] how to reinterpret this kind of dynamics in matrix terms. Given $\mathbf{x} \in \mathbb{R}^n$, we define an extended state vector $\bar{x}$ as

$$\bar{x} = \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \tag{1}$$

[1] In the literature one often finds also the equivalent term *piecewise affine*.

Similarly, given an $n \times n$ matrix A and a vector $\mathbf{b} \in \mathbb{R}^n$ we define $\bar{A}$ to be the $(n+1) \times (n+1)$ block matrix

$$\bar{A} = \begin{pmatrix} A & \mathbf{b} \\ 0_{1\times n} & 0 \end{pmatrix} \tag{2}$$

In this way the affine evolution equation $\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{b}$ can be formulated more compactly as $\dot{\bar{x}} = \bar{A}\bar{x}$.

In this paper we are going to study the stability of a certain class of PWL switched systems using *piecewise-linear Lyapunov functions*, as described for instance in [Joh03], Section 4.10. The advantage of piecewise-linear functions over piecewise-quadratic ones is that the former are tipically the outcome of a *linear programming* (LP) problem, whereas piecewise-quadratic Lyapunov functions are usually computed by solving a system of *linear matrix inequalities* (LMI), which is computationally more difficult (although faster methods are sometimes available, see e.g. [AGS$^+$16]).

A difference with respect to the approach in [Joh03] is that we do not require our Lyapunov functions to be continuous on the switching surfaces. (In particular, this means that we can do without the complex machinery of "continuity matrices".) We shall only require that V decreases on every switch, possibly in a discontinuous manner. The reasons for this choice will become clear later on.

We now briefly describe the model introduced in [AF13] to describe the resilience of interconnected systems.

The starting point is a set of n *linear* dynamical systems which are interconnected according to the edges of a directed graph G. The state of each system $k \in \{1, \dots, n\}$ is abstracted into a single variable $x_k \in [0,1] \subseteq \mathbb{R}$ ("percentage of service loss"), with 0 indicating regular functioning and 1 indicating total failure.

Each system has a corresponding *service recovery rate* $\mu_k > 0$, which quantifies its recovery capability when the state is perturbed away from zero. Moreover, for each pair of systems (i,k) connected by an arc $i \to k$ in the graph G the following parameters are defined:

- a *coupling coefficient* $\alpha_{ik} > 0$ between the state variables of the two systems;
- a *state threshold* $\sigma_{ik} \in (0,1)$ representing the maximum percentage of service loss beyond which the system i is considered to be failed for the descendant system k;
- a *service loss rate* $\lambda_{ik} > 0$ representing the rate of approaching failure of system k when the ancestor system i is malfunctioning.

Let I_k be the set of incoming neighbours of node k. System k is in *nominal operation mode* when $x_i \leq \sigma_{ik}$ for every $i \in I_k$. The evolution equation for the system is then

$$\dot{x}_k = -\mu_k(x_k - \sum_{i \in I_k} \alpha_{ik} x_i) + d_k, \tag{3}$$

where d_k is a real function modeling the (time-varying) external disturbance experienced by the k-th system.

If any of the input dependencies are malfunctioning the systems enters a failure mode. Denoting by m_k the cardinality of I_k, there are $2^{m_k} - 1$ such modes for system k: m_k where a single ancestor system is malfunctioning, $\binom{m_k}{2}$ where two ancestors are malfunctioning, and so on. For each $i \in I_k$, if $x_i > \sigma_{ik}$ is the only failed input then the evolution equation becomes

$$\dot{x}_k = \lambda_{ik}(1 - x_k), \tag{4}$$

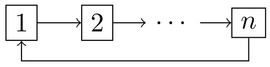
and more generally if $J_k \subseteq I_k$ is the subset of failed inputs we put[2]

$$\dot{x}_k = \sum_{j \in J_k} \lambda_{jk}(1 - x_k). \tag{5}$$

This system of evolution laws defines a piecewise-linear switching system whose state space is the n-dimensional hypercube $[0,1]^n \subset \mathbb{R}^n$. The basin of attraction of the equilibrium point located at the origin will be called the *resilience region* of the system.

III. SYSTEM CONFIGURATION

In the sequel we shall only consider the models that correspond to the family of graphs G_n defined as follows. For each $n \in \mathbb{N}$, the graph G_n has n nodes; for each $i = 1 \dots n-1$, node i node is connected by an edge to node $i+1$; finally, node n is connected to node 1. In other words, we consider a chain of systems with a final "feedback" connection:



For $n = 2$, this is exactly the model that was considered originally in [AF13]. As in the latter paper, we are only interested in evaluating the resilience against "instantaneous" disturbances that shift the state of the system away from the equilibrium, so from now on we shall assume that the noise term d_k in the evolution equation (3) is zero.

To each edge[3] $i \to i+1$ in the graph G_n there corresponds a threshold $\sigma_{i,i+1} \in [0,1]$ that we are going to denote more briefly by σ_i. This notation is quite natural in this setting, since the value σ_i pertains to the state variable of the i-th subsystem.

Each threshold divides the segment $[0,1]$ along the i-th axis of the state space in two parts, $[0,1] = [0,\sigma_i] \cup [\sigma_i, 1]$. It follows that the state space is partitioned into a total of 2^n cells. Each cell is a n-dimensional (rectangular) hypercuboid; the switching surfaces are the hypersurfaces $x_i = \sigma_i$ for $i = 1 \dots n$.

The remaining parameters defining the systems are:

- n service recovery rates $\mu_1, \dots, \mu_n$;
- n service loss rates $\lambda_1, \dots, \lambda_n$, where (compared to the notations of Section II) λ_i is a shorthand for $\lambda_{i-1,i}$;
- n coupling coefficients $\alpha_1, \dots, \alpha_n$, where α_i is a shorthand for $\alpha_{i-1,i}$.

[2] This rule for combining failures is not explicitly stated in [AF13].

[3] Here and in the sequel every index running from 1 to n is understood to be taken modulo n, so that $n+1 = 1$.

It is important to note that the model is *positive* (all state variables are constrained to assume only positive values) and *bounded*.

Let us introduce a more systematic labeling for the cells of this switching system. For any subset S of $\{1, \dots, n\}$ we shall denote by R_S the region where every system whose index belongs to S is over the threshold, and all the remaining systems are under the threshold. In particular $R_\emptyset$ denotes the region corresponding to the normal operation mode, i.e. the hypercuboid

$$R_\emptyset = [0, \sigma_1] \times \cdots \times [0, \sigma_n],$$

whereas for instance

$$R_{\{1\}} = [\sigma_1, 1] \times [0, \sigma_2] \times \cdots \times [0, \sigma_n]$$

represents the cell in which system 1 is failed for system 2, and so on, up to

$$R_{\{1 \dots n\}} = [\sigma_1, 1] \times \cdots \times [\sigma_n, 1]$$

which denotes the region of total failure.

The resilience region obviously includes the whole of $R_\emptyset$ and is disjoint from region $R_{\{1 \dots n\}}$. We are interested in delineating its boundary in the remaining $2^n - 2$ regions where some, but not all, of the subsystems are failing.

Let us reformulate the dynamics of the system in matrix form, using the linear embedding of affine dynamics described by equations (1–2). In region $R_\emptyset$ the system is linear with evolution matrix

$$A_\emptyset = \begin{pmatrix} -\mu_1 & 0 & 0 & \dots & \mu_1\alpha_1 \\ \mu_2\alpha_2 & -\mu_2 & 0 & \dots & 0 \\ 0 & \mu_3\alpha_3 & -\mu_3 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \mu_n\alpha_n & -\mu_n \end{pmatrix} \qquad (6)$$

An easy proof by induction shows that

$$\det A_\emptyset = (-1)^n \mu_1 \dots \mu_n (1 - \alpha_1 \dots \alpha_n)$$

It follows that the dynamics is nonsingular if and only if $\alpha_1 \dots \alpha_n \neq 1$; in the following we are going to suppose this is the case.

We also assume that in this region the system is stable. Using the well-known Routh-Hurwitz criterion for positive systems it is easy to check that this happens if and only if $\alpha_1 \dots \alpha_n < 1$.

When some system fails, for instance when system n fails for system 1, the dynamics switches from linear to affine and the new (extended) evolution matrix reads

$$\bar{A}_{\{1\}} = \begin{pmatrix} -\lambda_1 & 0 & 0 & \dots & 0 & \lambda_1 \\ \mu_2\alpha_2 & -\mu_2 & 0 & \dots & 0 & 0 \\ 0 & \mu_3\alpha_3 & -\mu_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & \mu_n\alpha_n & -\mu_n & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 \end{pmatrix}$$

Similarly, the failure of the i-th system causes the i-th row of the matrix (6) to be replaced with a row of the form

$$0 \quad \dots \quad 0 \quad -\lambda_1 \quad 0 \quad \dots \quad 0 \quad \lambda_i$$

In particular for the region number $2^n - 1$ where all subsystems have failed the evolution matrix is

$$\bar{A}_{\{1 \dots n\}} = \begin{pmatrix} -\lambda_1 & 0 & \dots & 0 & \lambda_1 \\ 0 & -\lambda_2 & 0 & \dots & \lambda_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -\lambda_n & \lambda_n \\ 0 & \dots & 0 & 0 & 0 \end{pmatrix}$$

It is interesting to note that these models are actually exactly solvable: the shape of the resilience region can (at least in principle) be determined analitically. This was done in [AF13] for the case $n = 2$ with no coupling between the two subsystems and no external disturbances. The boundary of the resilience region turns out to be the part of the square $[0, 1]^2$ which lies below the two curves

$$x_2 = 1 - (1 - \sigma_2) \left(\frac{x_1}{\sigma_1} \right)^{\lambda_2/\mu_1}$$

and

$$x_2 = \sigma_2 \left(\frac{1 - x_1}{1 - \sigma_1} \right)^{\mu_2/\lambda_1}$$

which meet (with a corner) at the point (σ_1, σ_2).

A similar analysis can be performed for a system with $n > 2$ nodes; however, the explicit description of the boundary gets more and more complex as the dimension grows. Notice also that the functions defining the boundary are transcendental, so they are quite expensive to compute with arbitrary precision.

For this reason it is advantageous to have a quickly computable (e.g. piecewise-linear) approximation of the boundary. In the next Section we turn to this goal.

IV. FROM STABILITY TO RESILIENCE

The general summary of our approach is the following. We look for a Lyapunov function whose sublevel sets approximate the basin of attraction of the equilibrium at zero. This function must be computationally cheap to work with; the simplest choice is to consider a piecewise-linear function on a polyhedral partition of the state space. The partition is computed starting from the cells of the switching system by successive refinements. Since the models we use are positive and have no limit cycles, we expect piecewise-linear Lyapunov functions to be a good fit.

Let us describe in detail our approach to construct a piecewise-linear Lyapunov function for the family of switching models described in Section III.

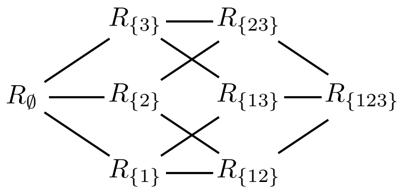
We start by looking for a piecewise-linear Lyapunov function defined on the same polyhedral partition of the state space given by the switching surfaces, namely the 2^n cells R_S for $S \subseteq \{1, \dots, n\}$. Actually, we already know that the cell corresponding to the choice $S = \{1, \dots, n\}$ is surely outside of the resilience region, so we can safely ignore this cell in what follows. Thus, we consider a Lyapunov function candidate of the form

$$V(x) = \begin{cases} \mathbf{k}_0 \cdot \mathbf{x} & \text{if } \mathbf{x} \in R_\emptyset, \\ \bar{k}_S \cdot \bar{x} = \mathbf{k}_S \cdot \mathbf{x} + k_{S,n+1} & \text{if } \mathbf{x} \in R_S. \end{cases} \tag{7}$$

Here $\mathbf{k}_0 \in \mathbb{R}^n$, $\bar{k}_S = (\mathbf{k}_S, k_{S,n+1}) \in \mathbb{R}^{n+1}$ and the index S runs over the proper nonempty subsets of the set $\{1, \ldots, n\}$. The total number of parameters to be determined is then $n + (2^n - 2)(n+1)$.

If we now require V to be continuous on *every* switching surface we see that as soon as $n > 2$ the continuity constraints eat up all the degrees of freedom in the multiple-failure regions. To clarify this point, let us start by noting that every cell in the decomposition $\{R_S\}$ can be put in one-to-one correspondence with a vertex of the hypercube $[0,1]^n$: for any subset $S \subseteq \{1, \ldots, n\}$ the region R_S corresponds to the vertex $V(S)$ whose coordinates are 1 for each axis x_i with $i \in S$ and 0 for each axis x_i with $i \notin S$.

The incidence relations between the vertices and the edges of an n-dimensional hypercube are described by the *hypercube graph* Q_n [HHW88]. Using the correspondence defined above, we can use this graph to describe the relationship between different regions in the partition $\{R_S\}$. For instance when $n = 3$ we have



The dimension of the boundary between two regions can be recovered from this graph: indeed $R_S \cap R_{S'}$ is a cuboid of dimension $n - k$, where k is the (geodesic) distance between R_S and $R_{S'}$ in the graph Q_n. For instance in the $n = 3$ case $R_\emptyset$ and $R_{\{1\}}$ intersect in a rectangle (dimension 2), $R_\emptyset$ and $R_{\{12\}}$ intersect in a segment (dimension 1) and finally $R_\emptyset$ and $R_{\{123\}}$ intersect in a single point (dimension 0).

We can now cook up a recipe to impose some reasonable continuity constraints on the candidate Lyapunov function (7). Our strategy is the following:

- for each $i \in \{1, \ldots, n\}$ we require $V_{\{i\}}$ to match $V_\emptyset$ on the whole $(n-1)$-dimensional cuboid $R_\emptyset \cap R_{\{i\}}$;
- for each $i < j \in \{1, \ldots, n\}$ we require $V_{\{ij\}}$ to match $V_\emptyset$ on the $(n-2)$-dimensional cuboid $R_\emptyset \cap R_{\{ij\}}$;
- for each $i < j < k \in \{1, \ldots, n\}$ we require $V_{\{ijk\}}$ to match $V_\emptyset$ on the $(n-3)$-dimensional cuboid $R_\emptyset \cap R_{\{ijk\}}$;

and so on. We remark that on each $(n-k)$-dimensional cuboid it is enough to select a subset of vertices forming a $(n-k)$-dimensional simplex, since two affine functions coinciding on (the vertices of) a simplex coincide everywhere.

The number of constraints generated by this rule is given by

$$\begin{aligned} n_c = \ & \tbinom{n}{1} n + \tbinom{n}{2}(n-1) + \cdots + \tbinom{n}{n-1} \cdot 2 \\ = \ & \textstyle\sum_{i=1}^{n-1} \tbinom{n}{i}(n-i+1). \end{aligned}$$

Using some well-known properties of the binomial coefficients we see that

$$n_c = (2^n - 2)(\frac{n}{2} + 1),$$

so that the number of free parameters in (7) is reduced to

$$n + (2^n - 2)(n+1) - n_c = 2^{n-1} n.$$

We shall also require that $V_\emptyset(\sigma_1, \ldots, \sigma_n) = 1$. This can be seen as a normalization condition for V; it is particularly convenient because the trajectories that hit the critical point $(\sigma_1, \ldots, \sigma_n)$ lie exactly on the boundary of the resilience region.

We now formulate the conditions which guarantee that the function (7) is a Lyapunov function for the system. We need two kinds of constraints:

- we should ensure that V is positive on each vertex of region $R_\emptyset$ and decreasing on every vertex in a boundary between two cells (it is of course sufficient to consider the vertices where no continuity conditions are imposed);
- we should ensure that V is strictly decreasing along the trajectories of the system in the interior of each cell. Since the dynamics there is linear, it suffices to check that

$$\begin{cases} \mathbf{k}_0 \cdot A_\emptyset \cdot \mathbf{v} < 0 & \text{for every vertex } \mathbf{v} \text{ of } R_\emptyset, \text{ and} \\ \bar{k}_S \cdot \bar{A}_S \cdot \bar{v} < 0 & \text{for every vertex } \bar{v} \text{ of } R_S. \end{cases}$$

If we group the entries of the vectors $\mathbf{k}_0$ and $\bar{k}_S$ in a single column matrix K, both these conditions can be expressed as a system of inequalities of the form $MK < 0$ for a suitable block matrix M.

Finally, we should select a suitable function of the coefficients of V to minimize in order to capture the largest possible subset of the resilience region. As V is everywhere positive, it suffices to minimize the sum of the values attained by V on a suitable set of $2^n - 1$ points (one for each region of the partition). Since (as explained above) each region is in one to one correspondence with a vertex of the cube $[0,1]^n$, it is natural to minimize the value of V on such points. Thus we define

$$\begin{aligned} F(K) := \ & V(1,0,\ldots,0) + V(0,1,0,\ldots,0) + \cdots + \\ & V(1,1,0,\ldots,0) + \ldots \end{aligned}$$

and look for the matrix $K = (\mathbf{k}_0, \bar{k}_{\{1\}}, \bar{k}_{\{2\}}, \ldots)$ which minimizes this function.

Summing up, we are led to the problem of minimizing $F(K)$ subject to the constraints:

$$\begin{cases} MK \leq 0 \\ EK = 0 \\ V(\sigma_1, \ldots, \sigma_n) = 1 \end{cases} \tag{8}$$

where E is a block matrix capturing the continuity constraints expressed earlier.

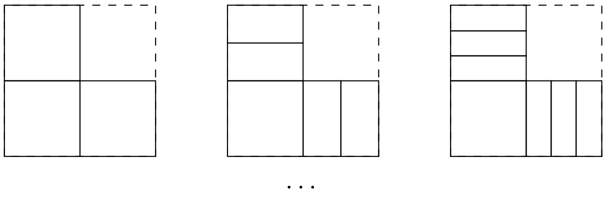
The solution of this linear programming problem (which we expect to be feasible for generic values of the parameters) determines a piecewise-linear Lyapunov function for the system. Then, by construction, the sublevel set

$$\{ x \in [0,1]^n \mid V(x) \le 1 \}$$

gives a piecewise-linear approximation to the resilience region.

The cell partition introduced above is rather coarse. As we saw in Section III, the boundary of the resilience region is given by a system of transcendental equations. In general, a single hyperplane will give a rather bad approximation for this boundary. It is then natural to seek a refinement of the partition $\{R_S\}$ in order to better capture the shape of the resilience region for the system.

For this purpose we can try to exploit again the special geometry of the state space of the model. The simplest idea is to consider refinements built according to the following rule, illustrated here for the $n = 2$ case:



In other words, we divide each single-failure region $R_{\{i\}}$ in $m > 1$ subregions and let the coefficients $\bar{k}_{\{i\}}$ in the expression (7) vary between those subregions.

Of course, things get quickly more complicated when $n > 2$: clearly, the subdivision described above automatically induces a subdivision of each double-failure region $R_{\{ij\}}$ in m^2 subregions, and so on. In general, one ends up with a total of

$$1+nm+\binom{n}{2}m^2+\cdots+\binom{n}{n-1}m^{n-1} = (m+1)^n - m^n$$

subregions. Notice that for $m = 1$ the above formula recovers the original number of regions $(2^n - 1)$.

The new candidate Lyapunov function now reads

$$V(x) = \begin{cases} \mathbf{k}_0 \cdot \mathbf{x} & \text{if } \mathbf{x} \in R_\emptyset, \\ \bar{k}_{S,j_S} \cdot \bar{x} & \text{if } \mathbf{x} \in R_{S,j_S} \end{cases} \tag{9}$$

where, as before, the index S runs over the proper nonempty subsets of the set $\{1, \ldots, n\}$ and the additional index j_S runs over the set $\{1, \ldots, m^{|S|}\}$, where $|S|$ denotes the cardinality of the set S. The total number of parameters to be determined is in this case $n + (n + 1)((m + 1)^n - m^n - 1)$.

Generalizing the problem (8) to the new situation is, at least in principle, straightforward; the only difficulty has to do with the bookkeeping needed to manage the various constraints. In particular the continuity constraints adopted above for the case $m = 1$ can be generalized in multiple ways to the case $m > 1$, giving rise to linear programming problems which are more or less constrained. We shall leave a detailed analysis of such issues to future work.

V. EXPERIMENTS

To test our approach we consider a model inspired by a real urban wastewater treatment facility which ensures depollution and dumping at sea of urban wastewater produced by domestic and economic activities in an international tourist area encompassing a marine reserve. The facility handles an estimated maximum of 36,000 people, roughly equivalent to a waste-water supply of 250 liters per person, per day. The plant is heavily automatized: all biological, chemical and mechanical processes are supervised by a control system connected through the Internet with a remote monitoring center. The plant consists of several compartments characterized by interconnected tanks, making it an ideal test bench for our framework. In particular, we focus on a series of three interconnected subsystems inside the facility: a balancing (BA) tank, a denitrification tank (DE) and a nitrification-oxidification (NO) tank. The BA tank receives sewage from previous compartments and feeds the DE tank with enough liquor to maintain the level of the NO tank at a desired level. The DE tank regulates its own level, dumping excess fluid to the NO tank. Liquor from the NO tank simply falls by gravity to other compartments for further treatment. As it can be observed in Figure 1, the topology of the interconnected systems respects the one considered in Section III. The input flow is a single-step function at $0.08\,m^3/s$, corresponding to the average inlet flow of the real facility (approximately $280\,m^3/h$). Gaussian white noise is added to the single-step function in order to simulate hourly variation of the inlet flow. The absolute value of the variation is within 20% of the average flow, consistently with data recorded at the facility[4].

For each tank $i \in \{1, 2, 3\}$ of Figure 1, the performance loss x_i is obtained from the corresponding output flow u_i through a *figure-of-merit function*, i.e., a mapping from the space of state variables to the space of performances. This is the standard way in which state variables are mapped to performance indicators — see, e.g., [HRM12] — and our specific mapping can be observed in Figure 2. Here we posit that an output flow of $0.08\,m^3/s$ corresponds to maximum functionality, whereas flows outside the interval $[0.06; 0.10]\,m^3/s$ correspond to total loss of functionality. The state threshold σ is placed at $x_i = 0.8$, corresponding to flows in the interval $[0.064; 0.096]\,m^3/s$, i.e., a variation of ±20% consistent with the actual input data. The choice of these parameters is motivated by the fact that the average input flow to the facility is $0.08m^3/s$ with a variation of ±20%. Outside this range, flows are considered to be abnormal, whereas variations above ±25% are considered total loss of functionality. In our experiments $\sigma_{ij} = \sigma$ for all $i, j \in \{1, 2, 3\}$ with $i \neq j$.

In order to analyze the resilience of the overall system according to the framework presented in Section II, we need to compute the values of the service loss rates,

[4] The complete model, as well as instructions to run it, are available on the companion web site of the paper at https://gitlab.sagelab.it/armtac/ifm2019companion.git

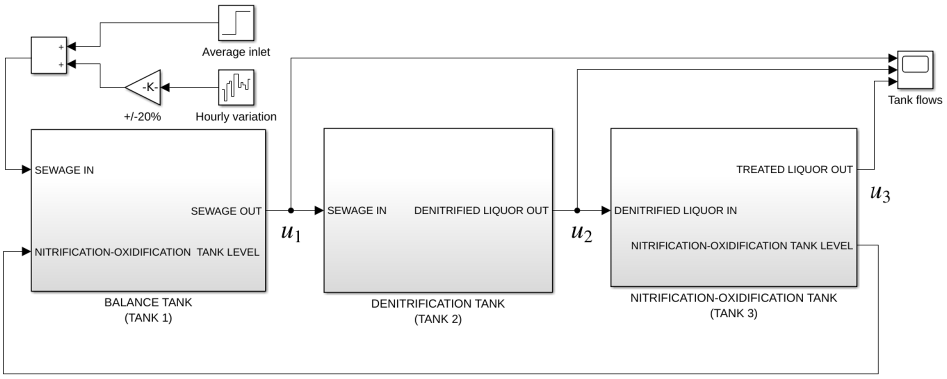


Fig. 1. Matlab/Simulink® model of three interconnected subsystems inside a wastewater treatment facility.

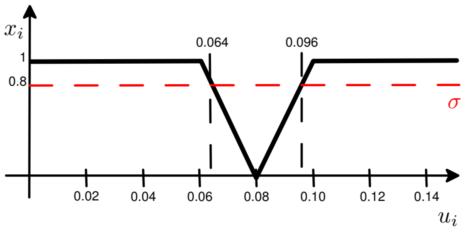


Fig. 2. Figure-of-merit (FOM) function to map tank outlet flow u_i to performance loss x_i. The red dotted line is the threshold σ beyond which the system is considered to be failed. Black dotted lines mark the points on the u_i axis within which the system is considered to be working properly.

TABLE I: System parameters computed by injecting faults in the system.

i	j	TTF	TTR	λ	μ	α
1	2	110.35	261.50	0.01458	0.00085	0.9
2	3	2251.00	1714.00	0.00071	0.00013	0.1
3	1	2127.00	1540.00	0.00076	0.00014	0.7

recovery rates and coupling coefficients. To this end, we consider the defintions presented in [AF13] and, given two interconnected systems $i, j \in \{1, 2, 3\}$ with $i \neq j$, we compute μ_{ij} and λ_{ij} as follows:

$$\mu_{ij} = -\frac{\ln(\sigma_{ij})}{TTR_{ij}} \qquad \lambda_{ij} = -\frac{\ln(1-\sigma_{ij})}{TTF_{ij}} \tag{10}$$

where

- TTR_{ij} (Time To Recover) is the time that system j takes to recover from a failure of system i and return within the state threshold σ_{ij} from the internal state $x_j = 1$, and
- TTF_{ij} (Time To Fail) is the time that system j takes to cross the state threshold σ_{ij} from the state $x_j = 0$.

Estimation of TTR, TTF and the coupling coefficient α can be done on the system model by introducing faults and then observing the behavior of relevant process variables. In particular we proceed as follows:

- To estimate TTF_{12} and TTR_{12} we introduce a stuck-at-0 fault in the pump that drains liquor from the BA tank into the DE tank. This causes the flow u_1 to decrease abruptly to 0. Consequently, the flow u_2 also decreases to 0 since the DE tank regulator tries to maintain a desired level in the tank, and stops sending fluid to the NO tank eventually. This effect is shown graphically in Figure 3 (top).
- To estimate TTF_{23} and TTR_{23} we introduce a stuck-at-0 fault in the pump that keeps the level of the DE tank stable. This causes the flow u_2 to decrease abruptly to 0. Consequently, the NO tank is not feeded anymore and starts emptying: reduction in the volume of the NO tank causes also reduction in its output flow which is due to gravity only.
- to estimate TTF_{31} and TTR_{31} we introduce a fault in the NO tank outlet by simulating a partial obstruction, which causes the NO tank to reduce its outlet flow below the average system input flow. Since the pump in the BA tank is regulated on the level of the NO tank, this will cause the output flow of the BA tank to decrease as well. The recovery from this fault is shown graphically in Figure 3 (bottom).

To estimate the coupling coefficient we proceed similarly, but instead of stuck-at faults, we consider drifting faults from nominal values and evaluate the ratio between the slopes of the overall flow variations. For instance, in order to estimate α_{12} we decrease the efficiency of the pump that drains liquor from the BA tank into the DE tank. Both u_1 and u_2 start to decrease until the normal functionality is resumed. The value of α_{12} is computed as the ratio between the slopes of u_1 and u_2 considering the onset of the failure as initial point and the end of the failure as final one. In all the cases above, faults are injected when the system reaches a regime condition (after 6 hours from the start) and last for a specified amount of time (1 hour).

Using Matlab we solved the linear programming problem (8) for the three-dimensional model with the coefficients listed in Table I. The complete matlab scripts and instructions on how to run them can be found on the companion web site mentioned above. We found the following values for the 27 parameters describing the Lyapunov function V:

$$\mathbf{k}_0 = \begin{pmatrix} 0.7081 \\ 0.0082 \\ 0.5338 \end{pmatrix} \bar{k}_{\{1\}} = \begin{pmatrix} 1.8903 \\ 0.0082 \\ 0.5338 \\ -0.9458 \end{pmatrix}$$

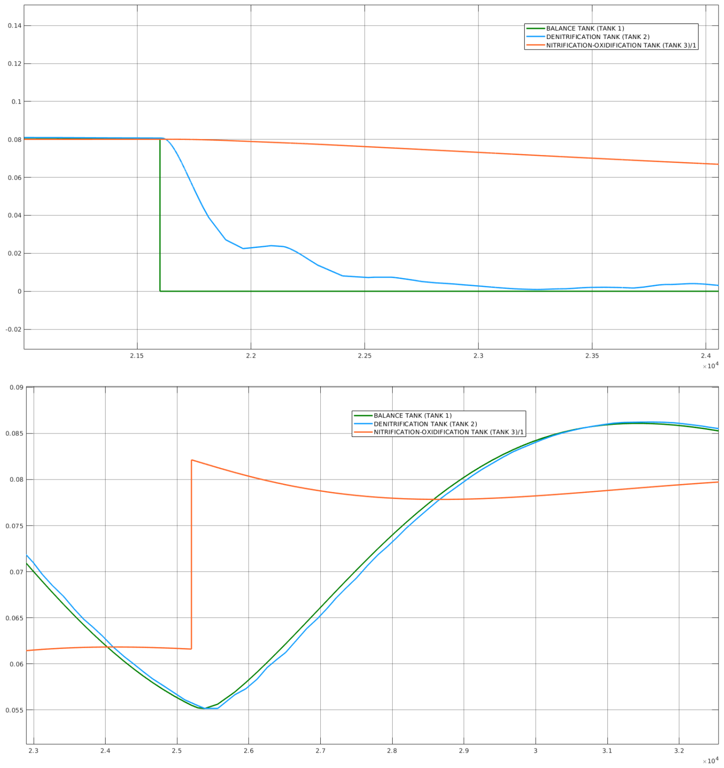


Fig. 3. Effects of injected faults on output flows. In each plot the lines represent the evolution of the tank outlet flows over time (x-axis in seconds, y-axis in cubic meters per second). The plot on the top shows the cascading effect of a "stuck-at-0" fault in the pump that drains liquor from the BA tank into the DE tank. The plot on the bottom shows the recovery from the effects of a partial obstruction in the NO tank on the flows of the BA and DE tanks.

$$\bar{k}_{\{2\}} = \begin{pmatrix} 0.7081 \\ 4.4070 \\ 0.5338 \\ -3.5191 \end{pmatrix} \bar{k}_{\{3\}} = \begin{pmatrix} 0.7081 \\ 0.0082 \\ 5.6899 \\ -4.1249 \end{pmatrix}$$

$$\bar{k}_{\{12\}} = \begin{pmatrix} 384.7189 \\ 4.4070 \\ 0.5338 \\ -310.7277 \end{pmatrix} \bar{k}_{\{13\}} = \begin{pmatrix} 1.8903 \\ 0.0082 \\ 5.6899 \\ -5.0707 \end{pmatrix}$$

$$\bar{k}_{\{23\}} = \begin{pmatrix} 0.7081 \\ 13.4647 \\ 5.6899 \\ -14.8901 \end{pmatrix}$$

Some graphs of the resulting piecewise-linear Lyapunov function are represented in figure 4 and figure 5 as slices in the (x_1, x_2) plane for various values of x_3.

We notice that the resulting Lyapunov function has a very steep slope in the region $\{12\}$, i.e., when system 3 is the only subsystem operating normally. This is presumably due to both the (comparatively) higher value of $\lambda_2 = \lambda_{12}$ and the strong coupling between the two failed subsystems. On the contrary, V is quite flat in the three single-failure regions and also in the region $R_{\{13\}}$, thanks to the good recovery capability of subsystem 1. Running the system model with faults injected confirms the picture obtained with the analysis.

VI. CONCLUSIONS

In this paper we have introduced a new method to compute piecewise linear Lyapunov functions to study the stability of interconnected systems, and thus the related resilience regions. We have shown that the method allows to analyze automatically the model of a real wastewater treatment facility by providing a computation approach and experimental results.

For a future development of this work it would be interesting to generalize our stability analysis to the full resilience model, with a nonzero time-varying external disturbance term in the evolution equation (3). Since the disturbance is not known a priori the model becomes a so-called *uncertain system* in this case, and different techniques need to be applied.

We would also like to consider interconnected systems with different (but still computationally tractable) topologies. In this regard, let us remark that the family of graphs described in Section III has the property that the number of interconnections grows only *linearly* as a function of the number of nodes n (as opposed to, e.g., a complete graph, where the number of edges grows quadratically in n). Other families of graphs with this property look relevant for the modeling of real-world systems.

In general, when some subsystem has more than one connection to other subsystems we expect to see more complex dynamical phenomena emerge (e.g. limit cycles). In these cases piecewise-linear Lyapunov functions are no longer a good fit, and it may become necessary to use a more general class of Lyapunov function candidates (piecewise quadratic, sum of squares, etc.).

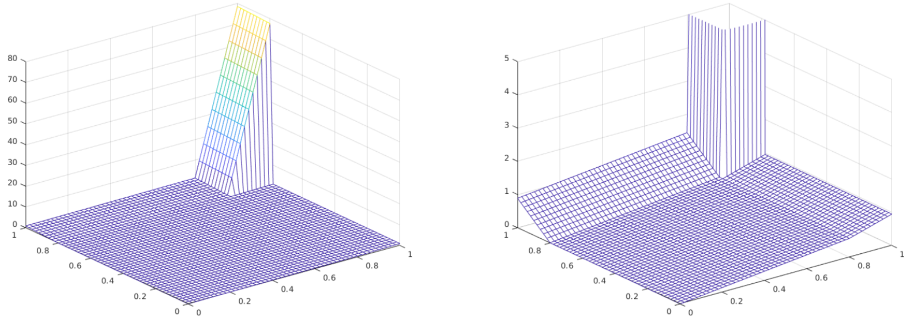


Fig. 4. Values of the Lyapunov function for $x_3 = 0$; the right-hand pane shows the same graph clipped to the region $V \leq 5$.

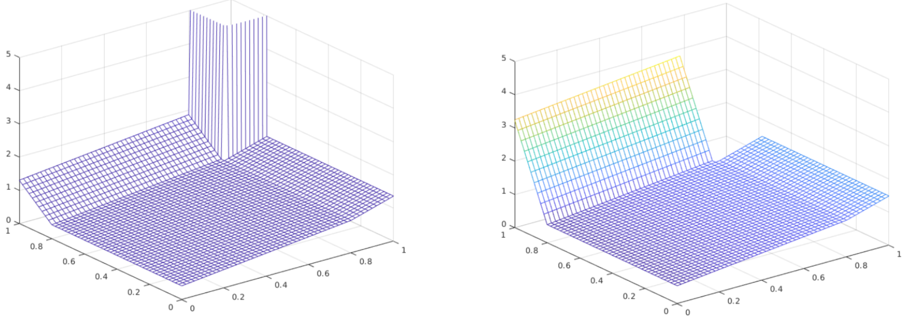


Fig. 5. Values of the Lyapunov function for $x_3 = 0.78$ and $x_3 = 0.82$ (threshold is $\sigma_3 = 0.8$).

References

[AF13] Angelo Alessandri and Roberto Filippini. Evaluation of resilience of interconnected systems based on stability analysis. In *Critical Information Infrastructures Security*, pages 180–190. Springer, 2013.

[AGS$^+$16] Xavier Allamigeon, Stéphane Gaubert, Nikolas Stott, Eric Goubault, and Sylvie Putot. A scalable algebraic method to infer quadratic invariants of switched systems. *ACM Trans. Embedded Comput. Syst.*, 15(4):69:1–69:20, 2016.

[BHSZ15] Fredrik Björck, Martin Henkel, Janis Stirna, and Jelena Zdravkovic. Cyber resilience–fundamentals for a definition. In *New contributions in information systems and technologies*, pages 311–316. Springer, 2015.

[CLM04] Paolo Crucitti, Vito Latora, and Massimo Marchiori. Model for cascading failures in complex networks. *Physical Review E*, 69(4):045104, 2004.

[GH15] Peter Giesl and Sigurdur Hafstein. Review on computational methods for Lyapunov functions. *Discrete Contin. Dyn. Syst. Ser. B*, 20(8):2291–2331, 2015.

[HHW88] Frank Harary, John P. Hayes, and Horng-Jyh Wu. A survey of the theory of hypercube graphs. *Comput. Math. Appl.*, 15(4):277–289, 1988.

[HRM12] Devanandham Henry and Jose Emmanuel Ramirez-Marquez. Generic metrics and quantitative approaches for system resilience as a function of time. *Reliability Engineering & System Safety*, 99:114–122, 2012.

[Joh03] Mikael Johansson. *Piecewise linear control systems*, volume 284 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 2003. A computational approach.

[Kha92] Hassan K. Khalil. *Nonlinear systems*. Macmillan Publishing Company, New York, 1992.

[LFZ17] X Liu, E Ferrario, and Enrico Zio. Resilience analysis framework for interconnected critical infrastructures. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 3(2):021001, 2017.

[Lib03] Daniel Liberzon. *Switching in systems and control*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 2003.

[Lou15] George Loukas. *Cyber-physical attacks: A growing invisible threat*. Butterworth-Heinemann, 2015.

[LPZ14] Xing Liu, Ionela Prodan, and Enrico Zio. On the resilience analysis of interconnected systems by a set-theoretic approach. *Safety and Reliability: Methodology and Applications, CRC Press, Leiden, The Netherlands*, pages 197–205, 2014.

[Oba13] Barack Obama. Presidential Policy Directive 21 (PPD21): Critical infrastructure security and resilience. *Washington, DC*, 2013.

[SG11] Zhendong Sun and Shuzhi Sam Ge. *Stability theory of switched dynamical systems*. Communications and Control Engineering Series. Springer, London, 2011.